

## Swarthmore College Works

---

Psychology Faculty Works

Psychology

---

2-1-2007

# School-Based Prevention Of Depressive Symptoms: A Randomized Controlled Study Of The Effectiveness And Specificity Of The Penn Resiliency Program

Jane Gillham

*Swarthmore College*, [jgillha1@swarthmore.edu](mailto:jgillha1@swarthmore.edu)

K. Reivich

D. R. Freres

T. M. Chaplin

A. J. Shatte

*See next page for additional authors*

Follow this and additional works at: <http://works.swarthmore.edu/fac-psychology>



Part of the [Psychology Commons](#)

---

### Recommended Citation

Jane Gillham, K. Reivich, D. R. Freres, T. M. Chaplin, A. J. Shatte, B. Samuels, A. G.L. Elkon, S. Litzinger, M. Lascher, R. Gallop, and M. E.P. Seligman. (2007). "School-Based Prevention Of Depressive Symptoms: A Randomized Controlled Study Of The Effectiveness And Specificity Of The Penn Resiliency Program". *Journal Of Consulting And Clinical Psychology*. Volume 75, Issue 1. 9-19.  
<http://works.swarthmore.edu/fac-psychology/72>

This Article is brought to you for free and open access by the Psychology at Works. It has been accepted for inclusion in Psychology Faculty Works by an authorized administrator of Works. For more information, please contact [myworks@swarthmore.edu](mailto:myworks@swarthmore.edu).

---

**Authors**

Jane Gillham, K. Reivich, D. R. Freres, T. M. Chaplin, A. J. Shatte, B. Samuels, A. G.L. Elkon, S. Litzinger, M. Lascher, R. Gallop, and M. E.P. Seligman



Published in final edited form as:

*J Consult Clin Psychol*. 2007 February ; 75(1): 9–19. doi:10.1037/0022-006X.75.1.9.

## School-Based Prevention of Depressive Symptoms: A Randomized Controlled Study of the Effectiveness and Specificity of the Penn Resiliency Program

**Jane E. Gillham,**

University of Pennsylvania and Swarthmore College

**Karen J. Reivich,**

University of Pennsylvania

**Derek R. Freres,**

University of Pennsylvania

**Tara M. Chaplin,**

University of Pennsylvania

**Andrew J. Shatté,**

University of Pennsylvania

**Barbra Samuels,**

University of Pennsylvania

**Andrea G. L. Elkon,**

University of Pennsylvania

**Samantha Litzinger,**

University of Pennsylvania

**Marisa Lascher,**

University of Pennsylvania

**Robert Gallop, and**

West Chester University

**Martin E. P. Seligman**

---

Copyright 2007 by the American Psychological Association

Correspondence concerning this article should be addressed to Jane E. Gillham, Psychology Department, Swarthmore College, 500 College Avenue, Swarthmore, PA 19081. [jgillha1@swarthmore.edu](mailto:jgillha1@swarthmore.edu) or [jgillham@psych.upenn.edu](mailto:jgillham@psych.upenn.edu).

Jane E. Gillham, Department of Psychology, University of Pennsylvania, and Psychology Department, Swarthmore College; Karen J. Reivich, Tara M. Chaplin, Andrew J. Shatté, Barbra Samuels, Andrea G. L. Elkon, Samantha Litzinger, Marisa Lascher, and Martin E. P. Seligman, Department of Psychology, University of Pennsylvania; Derek R. Freres, Annenberg School for Communication, University of Pennsylvania; Robert Gallop, Department of Mathematics Applied Statistics Program, West Chester University. Barbra Samuels is now at the Treatment Research Institute (Philadelphia). Andrea G. L. Elkon is now at the Georgia Regional Hospital at Atlanta (Atlanta, Georgia). Samantha Litzinger is now at the Department of Psychology, University of Tennessee. Marisa Lascher is now at Lehman Brothers (New York).

The Penn Resiliency Program for Children and Adolescents is owned by the University of Pennsylvania. The University of Pennsylvania has licensed this program to Adaptiv Learning Systems (King of Prussia, Pennsylvania). Andrew J. Shatté is employed part-time by Adaptiv Learning Systems, and Karen J. Reivich, Andrew J. Shatté, and Martin E. P. Seligman own stock in Adaptiv Learning Systems and could profit from the sale of this program. The other authors do not have financial relationships with Adaptiv Learning Systems.

University of Pennsylvania

## Abstract

The authors investigated the effectiveness and specificity of the Penn Resiliency Program (PRP; J. E. Gillham, L. H. Jaycox, K. J. Reivich, M. E. P. Seligman, & T. Silver, 1990), a cognitive-behavioral depression prevention program. Children ( $N = 697$ ) from 3 middle schools were randomly assigned to PRP, Control (CON), or the Penn Enhancement Program (PEP; K. J. Reivich, 1996; A. J. Shatté, 1997), an alternate intervention that controls for nonspecific intervention ingredients. Children's depressive symptoms were assessed through 3 years of follow-up. There was no intervention effect on average levels of depressive symptoms in the full sample. Findings varied by school. In 2 schools, PRP significantly reduced depressive symptoms across the follow-up relative to both CON and PEP. In the 3rd school, PRP did not prevent depressive symptoms. The authors discuss the findings in relation to previous research on PRP and the dissemination of prevention programs.

## Keywords

depression; prevention; children; adolescence

Several cognitive-behavioral interventions show promise in preventing depressive symptoms in youths (for recent reviews, see Horowitz & Garber, 2006; Merry, McDowell, Hetrick, Bir, & Muller, 2004). Among these are the Coping with Stress Course (Clarke & Lewinsohn, 1995), the Resourceful Adolescent Program (RAP; Shochet, Whitefield, & Holland, 1997), the LISA-T Program (Pössel, Horn, Groen, & Hautzinger, 2004), and the Penn Resiliency Program<sup>1</sup> (PRP; Gillham, Jaycox, Reivich, Seligman, & Silver, 1990). Despite this promise, however, depression prevention programs that demonstrate positive effects are rarely incorporated into school or clinical settings, and little is known about the effectiveness of most programs when delivered by the individuals who work in these settings.

In addition, when these programs do work, it is usually unclear whether the cognitive-behavioral therapy skills or other nonspecific factors (e.g., time spent in a structured after-school activity, attention from an adult, support from peers) are responsible. The majority of depression prevention studies compare cognitive-behavioral interventions with a no-intervention control. Only a few studies have compared prevention programs with attention control groups or alternate interventions. Pattison and Lynd-Stevenson (2001) found that PRP did not significantly reduce or prevent depressive symptoms relative to both no-intervention and attention control groups. The small sample size ( $N = 63$ ) may have limited power to detect effects, however. Merry, McDowell, Wild, Bir, and Cunliffe (2004) compared RAP with a placebo control group consisting primarily of group arts and crafts activities and found some support for the efficacy of RAP relative to placebo, but the effect size was small, and effects were not found across different measures of depressive symptoms. The placebo condition used by Merry et al. controlled for several factors that are

<sup>1</sup>PRP materials are available for use in research. Requests for the curriculum should be made to [info@pennproject.org](mailto:info@pennproject.org).

not specific to cognitive-behavioral interventions, including adult attention and opportunities to interact with peers. A stronger test of the cognitive-behavioral model would control for additional nonspecific factors, such as the discussion of stressors and topics relevant to depression.

Shatté (1997) and Reivich (1996) took this approach in their studies, which compared PRP with an alternate intervention (the Penn Enhancement Program [PEP]) and a no-intervention control. A total of 152 middle school students were randomly assigned to one of the three conditions. Group leaders were school teachers, school counselors, and psychology graduate students not affiliated with the research team. PRP and PEP both reduced depressive symptoms relative to the no-intervention control. PRP and PEP participants' depression scores were not significantly different at most of the follow-up assessments. The lack of differences may reflect limited power, as very large cell sizes are often needed to detect differences between active interventions. In addition, it is possible that differences would have emerged had the researchers followed participants for more than 1 year postintervention.

In the present study, we evaluated PRP's effectiveness as compared with PEP and a no-intervention control in a large sample (baseline  $N = 697$ ) over a 3-year follow-up period. Our major goals were as follows: (a) to investigate the effectiveness of PRP when delivered in schools by school teachers, school counselors, and other group leaders not affiliated with a research team, and (b) to evaluate intervention specificity. We examined two conceptualizations of depression prevention: (a) reduced levels of symptoms across an extended period of time (Jaycox, Reivich, Gillham, & Seligman, 1994), and (b) reduced onset of symptoms above a certain threshold (Gillham, Shatté, & Freres, 2000).

## Method

### Participants

This study was approved by the Institutional Review Board and by school administrators and school boards in each of the participating school districts. Figure 1 shows the participant flow from recruitment through the 3-year follow-up. Starting in 1997, two consecutive cohorts of students participated in this study. Each year, the research team sent letters and consent forms describing the project to parents of students in three middle schools (henceforth referred to as Schools A, B, and C) in a suburban metropolitan area in the United States. Recruitment materials described the study as an investigation of two interventions designed to help students cope with day-to-day stressors that are common in adolescence. We informed parents that those children who reported elevated symptoms at baseline would first be offered spots in the project and that other children would be included as space allowed. In total, parents of approximately 4,000 children received recruitment materials. Because the response rate was lower than anticipated, we decided to evaluate PRP with all children for whom we received consent as long as they were not suffering from a depressive disorder at baseline.

A total of 718 children and their parents consented to participate in the project. Response rates ranged from 15% to 22% across the three schools. These children completed baseline

measures, including the Children's Depression Inventory (CDI; Kovacs, 2001). The 175 children with baseline CDI  $\geq 13$  were administered the depressive disorders section of the Diagnostic Interview for Children and Adolescents (Reich, 2000). Eight children scored positive for major depressive disorder on the Diagnostic Interview for Children and Adolescents and were referred to therapy. Thirteen families dropped from the study after baseline and prior to randomization.

The remaining 697 children (376 boys and 321 girls) entered the intervention phase. Within cohort and school, we stratified children by grade, gender, and baseline CDI score (on the basis of a median split) and then used a computer-generated random numbers sequence to randomly assign participants to one of three study conditions: PRP ( $n = 232$ ), PEP ( $n = 231$ ), or Control (CON;  $n = 234$ ). Within each school and across all three schools combined, there were no significant between-conditions differences on baseline demographic characteristics or depressive symptoms.

At baseline, children's mean age was 12.13 years ( $SD = 1.03$ ). A total of 268 children (39%) were in sixth grade, 237 (34%) were in seventh grade, and 191 (27%) were in eighth grade (grade information was missing for 1 student). Age and grade did not significantly differ by school. Other demographic characteristics differed by school (see Table 1).

At baseline, mean levels of depressive symptoms ( $M = 8.45$ ,  $SD = 7.35$ , range = 0–45) were in the range typically found in community samples (Kovacs, 2001). Mean depressive symptoms differed by school,  $F(1, 694) = 9.52$ ,  $p < .01$ . Students in School A ( $M = 7.03$ ,  $SD = 7.41$ ) reported lower levels of depressive symptoms than those in School B ( $M = 9.51$ ,  $SD = 7.11$ ) or School C ( $M = 8.71$ ,  $SD = 7.34$ ), both  $ps < .01$ , whereas children in Schools B and C did not differ on baseline depressive symptoms.

### Intervention Conditions

PRP (Gillham et al., 1990) is a group intervention that teaches cognitive-behavioral and social problem-solving skills. PRP strives to teach students to think flexibly and accurately about the challenges and problems that they confront. Students learn about (a) the link between beliefs, feelings, and behaviors; (b) cognitive styles, including pessimistic explanatory styles; and (c) cognitive restructuring skills, including how to challenge negative thinking by evaluating the accuracy of beliefs and generating alternative interpretations. Students also learn a variety of techniques for coping and problem-solving, including assertiveness, negotiation, decision making, and relaxation. Students apply the cognitive and problem-solving techniques in their lives through group discussions and weekly homework assignments.

PEP (Reivich, 1996; Shatté, 1997) is a group intervention that focuses on many of the stressors associated with adolescent depression. Topics include peer pressure, ethical dilemmas, trust and betrayal, improving communication, friendships, family conflict, setting and achieving goals, self-esteem, and body image. Each session includes structured activities, role-playing, and guided discussion. PEP was designed to control for adult attention, group cohesion, social support, the discussion of depression-relevant topics,

sharing thoughts and feelings with peers, and engagement in homework and in-session activities. A detailed description of PEP is available from the authors on request.

PRP and PEP groups met after school once each week for a total of twelve 90-min sessions. Group leaders were school teachers, school counselors, and graduate students in school psychology, education, and clinical psychology (not affiliated with the research team) who participated in a 30-hr training workshop and biweekly group supervision with the PRP and PEP developers. During the study, 19 PRP groups and 19 PEP groups were implemented, each containing between 6 and 14 children. Teachers and counselors led 25 groups, and graduate students led 11 groups. Because of scheduling difficulties and a shortage of group leaders, research team members led one PRP group and one PEP group. No difference was found in effects for school staff (teachers and counselors) versus graduate students and researchers. Parents of children in PRP and PEP received \$5 per session that their child attended and a \$15 bonus if their child attended at least eight sessions.

In meetings and correspondence with group leaders and potential participants, we described both PRP and PEP as interventions that aim to increase well-being and help to buffer students against the effects of common stressors. We informed leaders and parents that the programs use different strategies to achieve this goal: PRP teaches specific coping and problem-solving skills, whereas PEP is more discussion oriented.

Students in the CON condition did not participate in PRP or PEP but completed the same assessments as students in the active interventions. In all three conditions, students and their families were permitted to seek counseling or therapy.

### Intervention Integrity and Attendance

Intervention sessions were audio-recorded. Following the intervention phase, we selected four lessons from each group for intervention adherence coding. We chose these lessons because they contained key content and included early, middle, and later sessions. Group leaders were not informed which lessons would be rated. Two research assistants rated each lesson on the presentation of 11–14 concepts, skills, or activities. For each content item, raters used a 7-point scale ranging from 1 (*no coverage*) to 4 (*satisfactory coverage*) to 7 (*excellent coverage*). Interrater reliability for the 1–7 ratings was good (PRP intraclass correlation [ICC] = .65; PEP ICC = .77). We computed two integrity scores for each group: (a) the degree covered score, calculated by averaging both raters' integrity item scores for each session and then averaging the session averages, and (b) the percentage covered score, calculated by averaging the percentage of items covered satisfactorily (4) across the sessions. Degree covered and percentage covered scores were highly correlated ( $r = .83$ ,  $p < .01$ ). Integrity scores could not be computed for three PRP and two PEP groups because leaders forgot to tape sessions or experienced equipment malfunctions.

Average integrity scores for PRP groups were 4.9 ( $SD = 0.48$ ) for degree covered and 80% ( $SD = 7.5\%$ ) for percentage covered. Average scores for PEP groups were 4.4 ( $SD = 0.36$ ) for degree covered and 68% ( $SD = 5.7\%$ ) for percentage covered. Integrity scores were higher for PRP than PEP groups, likely reflecting the increased structure of the PRP

curriculum. Mean integrity scores for PRP and PEP groups did not differ significantly by school.

On average, the 232 children assigned to PRP attended 6.71 ( $SD = 4.22$ ) sessions, and the 231 children assigned to PEP attended 7.11 ( $SD = 4.43$ ) sessions,  $t(461) = -1.01$ ,  $ns$ . Thirty-seven children (16%) assigned to PRP and 35 (15%) assigned to PEP did not attend any sessions. Of these, 27 in each condition completed at least one follow-up assessment, and their data were included in analyses. PRP attendance differed by school,  $F(1, 231) = 5.58$ ,  $p < .01$ . Attendance was higher in School A ( $M = 7.30$ ,  $SD = 3.97$ ) and School C ( $M = 7.38$ ,  $SD = 4.04$ ) than School B ( $M = 5.40$ ,  $SD = 4.40$ ),  $ps < .05$ . PRP attendance was not significantly different in Schools A and C. PEP attendance did not differ significantly by school.

### Assessment Schedule and Depressive Symptom Measures

Children in all three conditions completed questionnaires at baseline and at follow-up assessments 2 weeks postintervention and every 6 months thereafter for 3 years. Families received \$12 for each assessment.

This article focuses on results for the study's primary outcome variable: depressive symptoms, as assessed with the CDI and the Children's Depression Rating Scale—Revised (CDRS–R; Poznanski & Mokros, 1996). The CDI is a widely used measure of depressive symptoms in children and adolescents, which has adequate reliability and validity (Kovacs, 2001). We used recommended cutpoints of CDI greater than or equal to 13 to categorize depressive symptoms as at least mild in intensity (referred to as *elevated symptoms* within), and greater than or equal to 19 to denote depressive symptoms as at least moderate in intensity (referred to as *high symptoms* within). These cutpoints are near the 70th and 90th percentiles, respectively, for this age group on the basis of the CDI standardization norms (Kovacs, 2001).

We also examined clinical symptoms with the CDRS–R. At each follow-up assessment, we invited children with a CDI score greater than or equal to 13 to participate in the CDRS–R, a semistructured interview assessing symptoms of depression in children and adolescents (Poznanski & Mokros, 1996). Trained research assistants conducted the interviews, and a second research assistant coded the interview tape for reliability purposes. Interviewers and coders were not informed of participants' condition assignments. Interrater reliability was good ( $ICC = .79$ ). A CDRS–R  $t$  score greater than or equal to 65 identified children likely to be suffering from a depressive disorder. Clinical psychologists working on the project contacted parents of these children to inform them of the assessment results and offer referrals. For analytic purposes, children with CDI scores less than 13 were assumed to have a CDRS–R  $t$  score less than 65.

During the 3-year follow-up, 235 (36%) of the 651 children who completed follow-up assessments scored greater than or equal to 13 on the CDI and were invited to interviews at one or more follow-up points. Of these 235 children, 62 (20 from PRP, 21 from PEP, and 21 controls) did not complete CDRS–R interviews and so were excluded from the CDRS–R analyses. Thus, CDRS–R analyses used data from 589 children: the 173 children with CDI



scores greater than or equal to 13 who completed the interviews and the 416 children who consistently scored less than 13 on the CDI during the follow-up.

### Data Analytic Strategy

**Power analyses**—Our goal was to detect small-to-medium intervention effects ( $d = 0.30$ ) allowing for 20% attrition. We planned an initial cell size greater than or equal to 220 (final cell size of  $n = 176$ ) to achieve a power of greater than or equal to .80 to detect a difference at  $\alpha = .05$  (two-tailed). Attrition was higher than expected, however, and our final sample size at the 3-year follow-up assessment allowed us to detect effect sizes greater than or equal to .45 at a power of greater than or equal to .80 and  $\alpha = .05$  (two-tailed).

**Inspection and transformation of data**—We evaluated assumptions of normality of continuous variables by examining histograms and calculating skewness and kurtosis scores at each assessment point. CDI scores were skewed at several assessment points so square root transformations were used for these data. Table 2 shows raw means to ease interpretation.

**Attrition and baseline differences**—We used analysis of variance (ANOVA) and chi-square analyses to compare baseline depressive symptoms and demographic characteristics of children who did and did not complete any follow-up assessments. The possibility of differential attrition was examined by testing for between-conditions differences in these variables among children who did not complete any follow-up assessments, as well as among children who did not complete the final (3-year) follow-up assessment.

We used ANOVAs and chi-square analyses to compare baseline depressive symptoms and demographic characteristics across the three study conditions. These analyses tested for differences among all participants who completed baseline (*baseline sample*) and among participants for whom we had follow-up data (*follow-up sample*).

**Intervention effects**—We evaluated intervention effects on depressive symptoms using intent-to-treat analyses that included data from all participants selected into the study who completed the baseline assessment and any of the follow-up assessments (the follow-up sample). We operationalized prevention in two ways, as (a) prevention of symptoms over time, or lower levels of depressive symptoms relative to the comparison group over the follow-up period, and (b) prevention of onset of symptoms above specified levels of severity: CDI scores greater than or equal to 13 or 19, and CDRS-R  $t$  scores greater than or equal to 65.

Mixed-model (MM) ANOVAs tested prevention of symptoms over time. These analyses predicted depression scores over time from condition with baseline depression scores covaried. Additional analyses examined potential school effects by predicting depression scores over time from condition, school, and the Condition  $\times$  School interaction with baseline scores covaried. Examination of the model fit statistics, such as the Akaike information criterion and the Bayesian information criterion, as well as the likelihood ratio tests, for several covariance structures indicated that the unstructured covariance structure was appropriate. When the MM ANOVA analyses revealed a significant or marginally

significant main effect of condition,  $t$  tests (with test statements) compared the estimated marginal means for each pair of conditions from the overall MM ANOVA. We also conducted analyses of covariance (ANCOVAs) for each assessment (predicting depressive symptoms at the given assessment from condition, with baseline depression score covaried) to enable comparisons with prevention studies that report effects at each assessment point.

We calculated effect sizes based on Cohen's  $d$  using the difference (PRP minus CON, PRP minus PEP) in estimated marginal means. For MM ANOVAs, we divided the difference in estimated marginal means by the square root of the covariance estimate following recommendations from Raudenbush and Xiao-Feng (2001). For ANCOVAs, we divided the difference in estimated marginal means by the pooled standard deviation. Negative effect sizes indicate that PRP group depression scores were lower than CON (or PEP) scores over the follow-up period. For PEP versus CON comparisons, negative effect sizes indicate lower scores in PEP than CON. We used two-tailed alphas. We report significant effects and nonsignificant trends ( $p < .10$ , referred to as *tendencies* within).

Chi-square analyses tested prevention of onset of elevated, high, and clinical levels of symptoms. Analyses examining prevention of elevated and high levels of symptoms used data from participants who scored below the given cutpoint (CDI < 13 and CDI < 19, respectively) at baseline.

**Hypotheses**—We predicted that PRP would prevent depressive symptoms relative to CON and PEP. Specifically, we expected that during the follow-up, PRP participants would (a) report lower mean levels of depressive symptoms and (b) show lower rates of onset of elevated, high, or clinical levels of symptoms than participants in the other conditions.

## Results

### Overall

**Attrition and baseline differences**—Of the 697 children who completed baseline, 46 did not complete any follow-up assessments. These children were evenly distributed across the study conditions and represented between 6% and 7% of the children in each condition. Children who did not complete any follow-ups were older than the children who completed follow-ups,  $F(1, 663) = 6.99, p < .01$ . We found no significant between-conditions differences in depressive symptoms or demographic characteristics among children who did not complete any follow-up assessments.

A total of 396 participants did not complete the final 3-year follow-up assessment. These children were evenly distributed across the study conditions and represented between 56% and 59% of the baseline samples in each condition. Children who did not complete the 3-year follow-up assessment were older, came from households with lower incomes, and were less likely to have parents who were married than children who completed this assessment,  $F(1, 662) = 5.01, p < .05$ ;  $F(1, 438) = 5.83, p < .05$ ; and  $\chi^2(1, N = 672) = 10.28, p < .01$ . We found no significant between-conditions differences in depressive symptoms or demographic characteristics among children who did not complete the final 3-year follow-up.

We found no significant differences by condition on baseline depressive symptoms or demographic variables in the baseline sample. In the follow-up sample, parents' marital status varied by condition,  $\chi^2(6, N = 643) = 13.03, p < .05$ , with PEP participants less likely than PRP participants or controls to have parents whose marital status was listed as "other." PRP and CON conditions did not differ on parental marital status. Analyses covarying parents' marital status are consistent with the outcome analyses reported within.

**Intervention effects**—PRP did not significantly reduce depressive symptoms over the follow-up period relative to CON,  $t(622) = -1.29, ns, d = -0.09$  (95% CI =  $-0.23, 0.05$ ); or PEP,  $t(623) = -0.31, ns, d = -0.02$  (95% CI =  $-0.16, 0.11$ ). PEP also did not reduce depressive symptoms significantly relative to CON,  $t(628) = -0.97, ns, d = -0.07$  (95% CI =  $-0.21, 0.07$ ). ANCOVAs comparing pairs of conditions at each assessment revealed no significant intervention effects and only one trend. PRP tended to prevent depressive symptoms relative to CON at the 12-month follow-up,  $F(1, 324) = 3.11, p < .10, d = -0.18$  (95% CI =  $-0.37, 0.02$ ).

PRP prevented elevated symptoms (onset of CDI  $\geq 13$ ) relative to CON,  $\chi^2(1, N = 325) = 5.29, p < .05$ , odds ratio = .61 (95% CI = 0.40, 0.95), but not PEP,  $\chi^2(1, N = 330) = 0.03, ns$ , odds ratio = .96 (95% CI = 0.60, 1.54). PEP prevented elevated symptoms relative to CON,  $\chi^2(1, N = 333) = 4.64, p < .05$ , odds ratio = .64 (95% CI = 0.42, 0.98). Among children with baseline CDI scores less than 13, 33 (20%) of 161 in PRP, 36 (21%) of 169 in PEP, and 52 (32%) of 164 in CON scored greater than or equal to 13 at least once over the follow-up. There were no significant intervention effects on high (onset of CDI  $\geq 19$ ) or clinical symptoms (CDRS-R  $t \geq 65$ ).

**School effects**—Intervention effects differed by school. A MM ANOVA predicting follow-up CDI scores from condition, school, and the School  $\times$  Condition interaction (with baseline symptoms covaried) revealed no main effect of condition, a significant main effect for school,  $F(1, 617) = 3.16, p < .05$ , and a significant School  $\times$  Condition interaction,  $F(4, 619) = 2.61, p < .05$ . Pairwise comparisons indicated that follow-up CDI scores were significantly worse in School C than School A ( $p < .05$ ), and tended to be worse in School B than School A ( $p < .10$ ). Follow-up CDI scores in School B and School C did not significantly differ.

To locate the source of the School  $\times$  Condition interaction, we conducted MM ANOVAs predicting CDI score from condition, school, and the School  $\times$  Condition interaction for each pair of schools. These analyses revealed no School  $\times$  Condition interaction for the School A and School B comparison but marginal and significant School  $\times$  Condition interactions for the comparisons of School C to Schools A and B,  $F(2, 419) = 2.89, p = .057$ , and  $F(2, 445) = 4.16, p < .05$ , respectively. An analysis predicting CDI from condition, School A + B versus C, and their interaction revealed no significant main effect of condition but a significant School (A + B vs. C)  $\times$  Condition interaction,  $F(2, 609) = 5.11, p < .01$ . To reduce the number of analyses and retain statistical power, we examined outcomes for Schools A and B (combined) and School C.

## Schools A and B

**Attrition and baseline differences**—We found no evidence for differential attrition in Schools A and B. One variable—parents' marital status—differed by condition in the baseline sample and in the follow-up sample,  $\chi^2(6, N = 411) = 15.02, p < .05$ , and  $\chi^2(6, N = 392) = 17.10, p < .01$ , respectively. Parents of PEP participants were more likely to be divorced and less likely to list marital status as “other” than parents of PRP participants or controls. Parents of PRP participants and controls did not differ on marital status. Analyses covarying marital status were consistent with those reported within.

**Intervention effects**—In Schools A and B, PRP significantly reduced depressive symptoms over the follow-up period relative to CON,  $t(636) = -2.59, p < .05, d = -0.24$  (95% CI =  $-0.42, -0.06$ ), and PEP,  $t(629) = -2.03, p < .05, d = -0.18$  (95% CI =  $-0.36, -0.01$ ). PEP did not significantly reduce depressive symptoms relative to CON,  $t(640) = -0.59, ns, d = -0.05$  (95% CI =  $-0.24, 0.13$ ). See Table 2 for means and ANCOVAs by assessment point.

PRP prevented elevated symptoms (onset of CDI  $\geq 13$ ) relative to CON,  $\chi^2(1, N = 205) = 8.03, p < .01$ , odds ratio = .38 (95% CI = 0.19, 0.75), but not relative to PEP,  $\chi^2(1, N = 212) = 1.38, ns$ , odds ratio = .66 (95% CI = 0.32, 1.33). PEP tended to prevent elevated symptoms relative to CON,  $\chi^2(1, N = 207) = 2.92, p < .10$ , odds ratio = .58 (95% CI = 0.31, 1.09). Among children with baseline CDI less than 13, 16 (15%) of 105 in PRP, 23 (22%) of 107 in PEP, and 32 (32%) of 100 controls scored greater than or equal to 13 on the CDI at least once during the follow-up (see Figure 2). There were no significant intervention effects on high symptoms. PRP tended to prevent clinical symptoms relative to CON,  $\chi^2(1, N = 237) = 2.74, p < .10$ , odds ratio = .51 (95% CI = 0.23, 1.14), but not PEP,  $\chi^2(1, N = 243) = 0.01, ns$ , odds ratio = .96 (95% CI = 0.40, 2.30). No difference was found between PEP and CON,  $\chi^2(1, N = 232) = 2.37, ns$ , odds ratio = .53 (95% CI = 0.24, 1.20). Across the follow-up period, 11 (9%) of 124 children in PRP, 11 (9%) of 119 in PEP, and 18 (16%) of 113 controls reported clinical symptoms on the CDRS-R.

## School C

**Attrition and baseline differences**—In School C, there were no between-conditions differences at baseline, and there was no evidence for differential attrition.

**Intervention effects**—PRP did not reduce depressive symptoms over the follow-up relative to CON,  $t(585) = 1.29, ns, d = 0.14$  (95% CI =  $-0.08, 0.36$ ). Contrary to our expectations, PEP significantly reduced depressive symptoms relative to PRP,  $t(594) = 2.08, p < .05, d = 0.24$  (95% CI = 0.01, 0.47). No difference was found between PEP and CON,  $t(592) = -0.86, ns, d = -0.10$  (95% CI =  $-0.32, 0.12$ ). See Table 2 for means and ANCOVAs by assessment. There were no significant intervention effects on elevated, high, or clinical symptoms (see Figure 2).

## Examining the School $\times$ Condition Interaction

We looked at the pattern of results on other measures (demographics, intervention integrity, attendance) to see whether any of these variables could explain the different effects across

schools. Group integrity scores did not differ by school. For most of the variables that differed by school, mean levels for School C were between the means of Schools A and B. One exception was race/ethnicity. Compared with Schools A and B, School C included a higher proportion of Caucasian participants and a lower proportion of participants who indicated their race was “other.” MM ANOVAs predicting follow-up CDI scores revealed no significant Race  $\times$  Condition interaction, however. Thus, none of the variables we examined explained the differential effects across schools.

## Discussion

The full sample findings provide little evidence for PRP’s effectiveness or specificity. PRP prevented the onset of elevated symptoms relative to a no-intervention control but not relative to PEP (an intervention that controlled for noncognitive–behavioral therapy specific factors). PRP did not reduce mean levels of depressive symptoms over the follow-up and did not prevent high or clinical levels of symptoms relative to either comparison group. PRP’s effects varied by school, however. In two schools, PRP reduced depressive symptoms over the follow-up relative to CON and PEP. In contrast, PRP was not effective in the third school.

In Schools A and B, PRP’s effects on depressive symptoms were long lasting. PRP participants reported fewer depressive symptoms than controls through the 30-month follow-up. Effect sizes were moderate at several assessment points. In addition, PRP prevented the onset of elevated symptoms and tended to prevent clinical levels of symptoms relative to no intervention control.

These findings indicate that school teachers and counselors can be trained to implement the intervention effectively and that schools are a viable setting in which to teach depression prevention skills. Given that depression is both very common and under-treated in youths (Hirschfeld et al., 1997), the importance of these findings should be underscored. Skills-based depression prevention programs, like PRP, may be an important addition to the educational goals of our school systems. By offering programs like PRP, schools may play a vital role in preventing a growing national health concern.

An important direction for future research is to determine whether PRP can prevent clinical levels of depressive symptoms when delivered by school teachers and counselors. The present findings provide more evidence for the prevention of mild-to-moderate depressive symptoms than for the prevention of clinical depression. This could reflect several factors, including (a) a limit of the PRP intervention, (b) the fact that most of our group leaders did not have clinical training and thus may have been unprepared to help the children at greatest risk, or (c) the low attendance of many PRP participants. The lack of findings for clinical depression also could be a function of statistical power as only a small proportion of children reported clinical levels of symptoms during the follow-up period.

### PRP’s Inconsistent Effects

Despite the positive findings in Schools A and B, PRP did not prevent depression in School C. We were unable to identify the source of the differential effects observed in the present

study. Group leaders in the three schools received similar training and supervision, and adherence to the PRP and PEP protocols did not vary significantly by school. Children's baseline depression scores, demographic characteristics, and attendance also did not account for the differential effects observed. We also are not aware of any differences in school services, curriculum, or in events that transpired within schools that could account for the discrepancy. Subtle differences among the schools in program endorsements, students' needs, or general school climate may have enhanced PRP's effectiveness in two schools and/or hindered its effectiveness in the third.

The inconsistency in PRP's effects within the present study mirrors its inconsistent effects across studies. PRP has reduced and prevented depressive symptoms relative to control in several studies (e.g., Gillham, Reivich, Jaycox, & Seligman, 1995; Yu & Seligman, 2002). For example, in the original evaluation of PRP, PRP halved the rate of moderate to severe symptoms. Effect sizes were moderate to large ( $d > 0.60$ ) 12–24 months postintervention (Gillham et al., 1995). However, other studies failed to replicate these findings (e.g., Pattison & Lynd-Stevenson, 2001; Roberts, Kane, Bishop, & Matthews, 2004). A review of PRP research suggests that PRP's effects are strongest when group leaders are the intervention developers, members of their research team, and graduate students who receive extensive training and supervision from the intervention developers. PRP's effects on depression are often smaller or nonsignificant when group leaders are school teachers, clinicians, or other researchers not affiliated with the PRP laboratory (Gillham, Hamilton, Freres, Patton, & Gallop, 2006). In other words, as research on PRP moves along the path from efficacy to real world implementation, intervention effects wane or become inconsistent. This pattern corresponds with recent findings for the RAP intervention (Harnett & Dadds, 2004) and with the observation that psychological interventions that yield substantial effects in research settings often produce small or no effects when delivered by community providers (Weisz, Donenberg, Han, & Weiss, 1995). A critical priority for prevention research is the development of effective dissemination strategies.

### Specificity

The School A and School B findings provide some support for specificity as mean depression scores were lower in PRP than PEP participants over the follow-up. Contrary to our expectations, however, PEP participants reported fewer depressive symptoms over the follow-up than PRP participants in School C. Because PRP was not beneficial relative to CON in School C, it is unclear whether PEP's superiority reflects the limited value of cognitive-behavioral skills or a more general failure of the PRP intervention in that school.

### Limitations

The study's limits include the low recruitment rate, substantial attrition, incomplete assessment of clinical depression, and the inclusion of graduate students as group leaders. In each school, fewer than 25% of students enrolled in the study. Enrollment rates were so much lower than we expected that we revised our original plan to select children with elevated symptoms and offered the program to all children who were not clinically depressed at the start of the study. The children who enrolled reported a wide range of symptoms at baseline, but we did not evaluate the interventions as delivered to most school



students in a school setting as in the recent study by Spence, Sheffield, and Donovan (2005). Our participants' demographic characteristics reflected the communities from which they came, but it is likely that these children (and their parents) differed in other important ways from those who did not sign up for the study. This limits the extent to which we can generalize from our findings.

Fewer than 50% of the original baseline sample completed the 3-year follow-up. Attrition was due to vicissitudes across schools but did not seem to vary by condition. For example, in School A, because of a change in administration, school staff stopped reminding children to attend assessment sessions. School B has had a high student attrition rate in general, with 10%–20% of students leaving the school each year. We tried to follow participants who moved through the mail, but compliance with these assessments was low, and we were often unable to find forwarding addresses. Attrition and noncompletion of assessments are common problems in longitudinal depression prevention studies. Attrition is often 40% or more in studies with 3 and 4 years of follow-up (e.g., Roberts et al., 2004; Spence et al., 2005). In our current work, we are implementing several strategies to minimize attrition, including increased incentives for complete assessments and strategies for tracking participants, such as requesting contact information for relatives.

Because of budget and staffing constraints, CDRS–R interviews were only conducted with children with elevated CDI scores (CDI = 13). CDI manual norms suggest that 8% of children who score below this cutpoint may have depressive disorders (Kovacs, 2001). In addition, only about 75% of children who were invited to interviews actually participated. Thus, we probably failed to detect clinical levels of symptoms in several children. During the study, we became concerned about another possible limitation of the triggered interview procedure—that it influenced questionnaire responses. Children may have become aware that elevated questionnaire scores triggered interviews, and children who wanted to avoid interviews may have endorsed few symptoms of depression on the CDI. It is striking that during a developmental phase when depression rates generally increase (Angold, Costello, & Worthman, 1998), depressive symptoms declined in all three conditions in our sample. In the end, the triggered interviews may have limited our ability to find intervention effects on both depression measures.

The inclusion of graduate students as group leaders is also somewhat problematic. The graduate students who led groups were not affiliated with our research laboratory and most planned to pursue careers as school psychologists or school counselors. Still, these graduate students may differ from the school personnel who would deliver an intervention like PRP once disseminated. An important finding is that analyses revealed no significant differences in intervention effects for graduate students versus school teachers and counselors.

Future research may also benefit from more sophisticated analyses of intervention implementation. Our intervention integrity scoring system weighted specific content items equally, but it is likely that some skills are more beneficial or require more explanation and discussion than others. An examination of the relationship between specific content covered and intervention outcome could yield valuable information about the mechanisms of PRP's effects and help set priorities for group leader training and supervision.

Our integrity coding system was designed to measure leaders' adherence to the intervention protocol and probably did not fully capture the variability in leaders' competence, or skill, in intervention delivery (McGlinchey & Dobson, 2003). Although competence and adherence are related, competent leaders may sacrifice adherence at times to allow for more in-depth discussion of group members' specific problems, review of concepts that seem confusing to group members, or more attention to group process. An assessment of competence could help to explain our differential findings; it may be that PRP group leaders in School C adhered to the manual but did not show true competence.

### Clinical Significance

We found little evidence for PRP's effectiveness or specificity in preventing depression in the full sample. However, in two of the three schools, PRP reduced depressive symptoms over a 2.5-year period and prevented elevated depressive symptoms. This suggests that PRP can have long-lasting effects when delivered by school teachers, counselors, and other providers in a school setting. Although PRP's effects are inconsistent, we believe that research that evaluates interventions in real world settings will lead to the development of dissemination procedures that enable schools to include powerful depression prevention programs as part of their educational mission.

### Acknowledgments

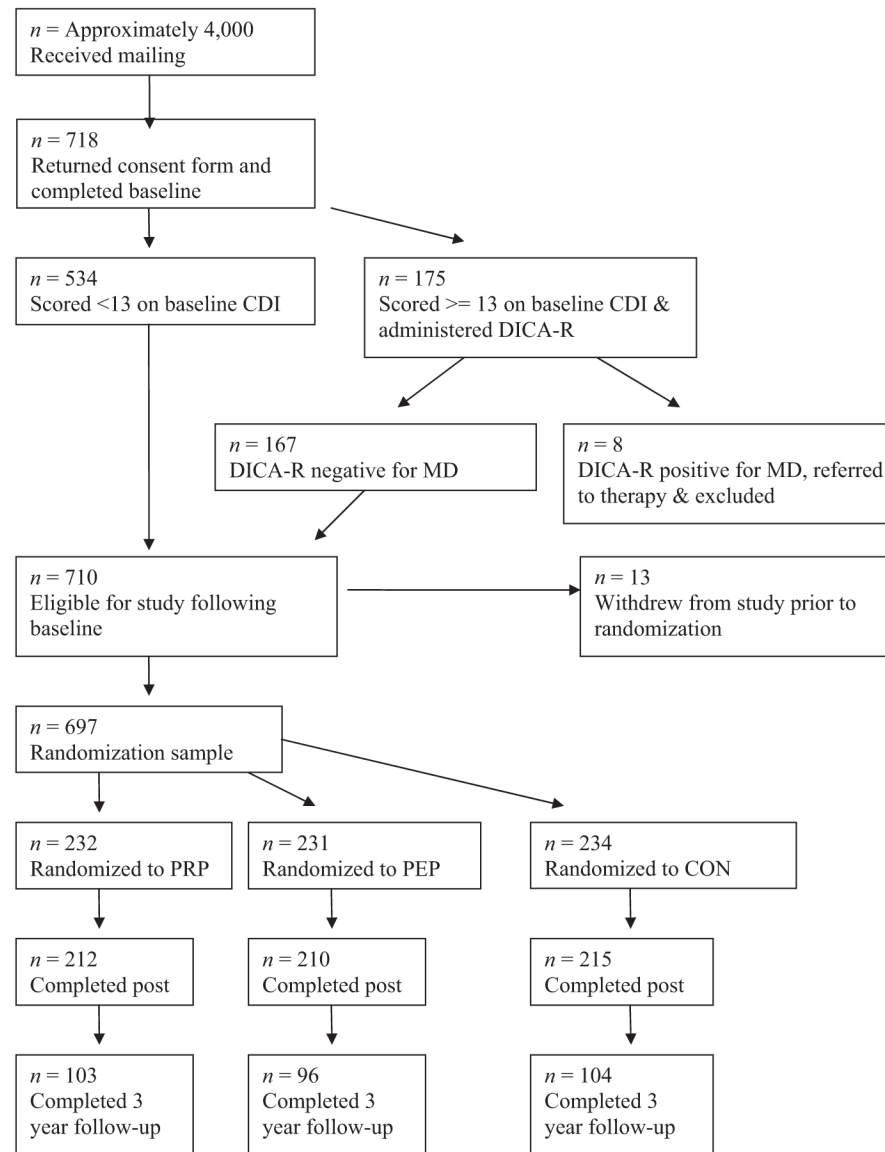
This research was funded by National Institute of Mental Health Grant MH52270. We are grateful to the children and parents who participated in the project, and to the school teachers, counselors, administrators, and personnel who assisted us in carrying out the project.

### References

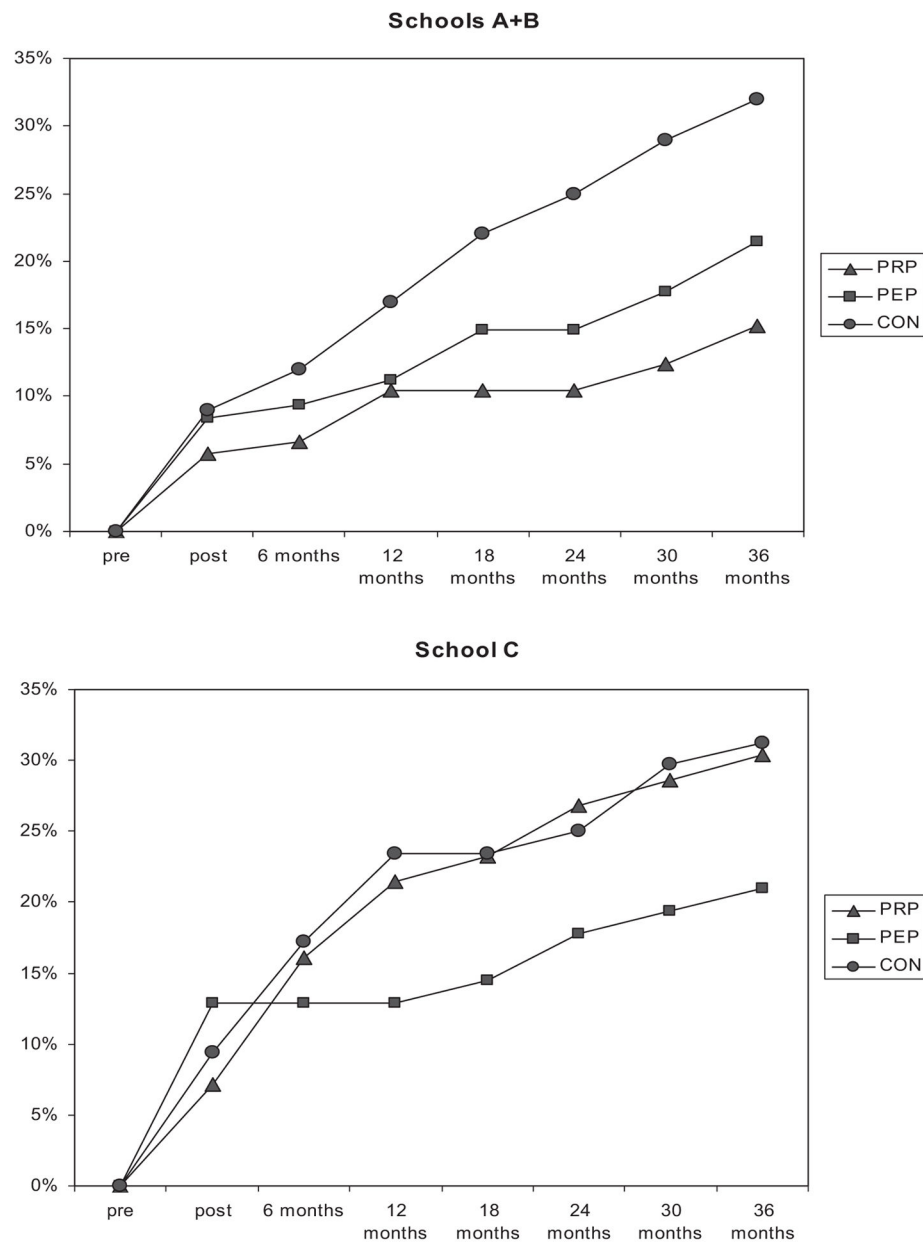
- Angold A, Costello EJ, Worthman CM. Puberty and depression: The roles of age, pubertal status, and pubertal timing. *Psychological Medicine*. 1998; 28:51–61. [PubMed: 9483683]
- Clarke, GN.; Lewinsohn, PM. Instructor's manual for the Adolescent Coping with Stress Course. 1995. Retrieved December 2, 2005, from Kaiser Permanente Center for Health Research Web site: <http://www.kpchr.org/acwd/acwd.html>
- Gillham JE, Hamilton J, Freres DR, Patton K, Gallop R. Preventing depression among early adolescents in the primary care setting: A randomized controlled study of the Penn Resiliency Program. *Journal of Abnormal Child Psychology*. 2006; 34:203–219. [PubMed: 16741684]
- Gillham, JE.; Jaycox, LH.; Reivich, KJ.; Seligman, MEP.; Silver, T. Unpublished manual. University of Pennsylvania; Philadelphia: 1990. The Penn Resiliency Program.
- Gillham JE, Reivich KJ, Jaycox LH, Seligman MEP. Preventing depressive symptoms in schoolchildren: Two year follow-up. *Psychological Science*. 1995; 6:343–351.
- Gillham JE, Shatté AJ, Freres DR. Depression prevention: A review of cognitive-behavioral and family interventions. *Applied & Preventive Psychology*. 2000; 9:63–88.
- Harnett PH, Dadds MR. Training school personnel to implement a universal school-based prevention of depression program under real-world conditions. *Journal of School Psychology*. 2004; 42:343–357.
- Hirschfeld R, Keller M, Panico S, Arons B, Barlow D, Davidoff F, et al. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. *Journal of the American Medical Association*. 1997; 277:333–340. [PubMed: 9002497]
- Horowitz JL, Garber J. The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *Journal of Consulting and Clinical Psychology*. 2006; 74:401–415. [PubMed: 16822098]



- Jaycox LH, Reivich KJ, Gillham J, Seligman MEP. Prevention of depressive symptoms in school children. *Behaviour Research & Therapy*. 1994; 32:801–816. [PubMed: 7993324]
- Kovacs, M. *Children's Depression Inventory manual*. North Tonawanda, NY: Multi-Health Systems; 2001.
- McGlinchey JG, Dobson KS. Treatment integrity concerns in cognitive therapy for depression. *Journal of Cognitive Psychotherapy*. 2003; 17:299–317.
- Merry S, McDowell H, Hetrick S, Bir J, Muller N. Psychological and/or educational interventions for the prevention of depression in children and adolescents. *Cochrane Database of Systematic Reviews*. 2004; 2:Article CD003380. Retrieved June 16, 2004, from The Cochrane Library Database.
- Merry S, McDowell H, Wild CJ, Bir J, Cunliffe R. A randomized placebo-controlled trial of a school-based depression prevention program. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2004; 43:538–547. [PubMed: 15100560]
- Pattison C, Lynd-Stevenson RM. The prevention of depressive symptoms in children: The immediate and long-term outcomes of a school based program. *Behaviour Change*. 2001; 18:92–102.
- Pössel P, Horn AB, Groen G, Hautzinger M. School-based prevention of depressive symptoms in adolescents: A 6-month follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2004; 43:1003–1010. [PubMed: 15266195]
- Poznanski, EO.; Mokros, HB. *Children's Depression Rating Scale—Revised (CDRS—R)*. Los Angeles: Western Psychological Service; 1996.
- Raudenbush SW, Xiao-Feng L. Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. *Psychological Methods*. 2001; 6:387–401. [PubMed: 11778679]
- Reich W. Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child and Adolescent Psychiatry*. 2000; 39:59–66. [PubMed: 10638068]
- Reivich, KJ. Unpublished doctoral dissertation. University of Pennsylvania; Philadelphia: 1996. The prevention of depressive symptoms in adolescents.
- Roberts C, Kane R, Bishop B, Matthews H. The prevention of depressive symptoms in rural children: A follow-up study. *International Journal of Mental Health Promotion*. 2004; 6:4–16.
- Shatté, AJ. Unpublished doctoral dissertation. University of Pennsylvania; Philadelphia: 1997. Prevention of depressive symptoms in adolescents: Issues of dissemination and mechanisms of change.
- Shochet, I.; Whitefield, K.; Holland, D. *Resourceful Adolescent Program: Group leader's manual*. Brisbane, Australia: Griffith University; 1997.
- Spence SH, Sheffield JK, Donovan CL. Long-term outcome of a school-based, universal approach to prevention of depression in adolescents. *Journal of Consulting and Clinical Psychology*. 2005; 73:160–167. [PubMed: 15709843]
- Weisz JR, Donenberg GR, Han SS, Weiss B. Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*. 1995; 63:688–701. [PubMed: 7593861]
- Yu DL, Seligman MEP. Preventing depressive symptoms in Chinese children. *Prevention and Treatment*. 2002; 5

**Figure 1.**

Flow diagram of recruitment, screening, and follow-up. CDI = Children's Depression Inventory; DICA-R = depressive disorders section of the Diagnostic Interview for Children and Adolescents; MD = major depressive disorder; PRP = Penn Resiliency Program; PEP = Penn Enhancement Program; CON = control; post = postintervention.



**Figure 2.**

Onset of elevated symptoms by school: Cumulative proportion of participants with a Children's Depression Inventory score greater than or equal to 13. PRP = Penn Resiliency Program; PEP = Penn Enhancement Program; CON = Control; pre = preintervention; post = postintervention.

Table 1

## Demographic Characteristics by School

Variable	School A ( <i>n</i> = 212)			School B ( <i>n</i> = 222)			School C ( <i>n</i> = 263)			Statistic
	<i>n</i>	%	<i>M</i>	<i>SD</i>	<i>n</i>	%	<i>M</i>	<i>SD</i>	<i>n</i>	%
Gender										
Boys	135	63			97	44			145	55
Girls	78	37			125	56			118	45
Race/ethnicity										
African American	12	6			40	18			10	4
Asian American	13	6			5	2			6	2
Caucasian	153	74			132	60			227	88
Latino/Latina	3	2			8	4			2	1
Other	25	12			35	16			13	5
Total with data	206				220				258	
Parents' marital status										
Married	141	73			75	34			134	51
Separated	1	1			15	7			10	4
Divorced	35	18			52	24			69	26
Other	16	8			76	35			48	18
Total with data	193				218				261	
Family income level (1997–1998)										
Mean level			4.49 <sub>a</sub>	1.54			2.58 <sub>c</sub>	1.22		
Level 1: <\$20,000	3	2			29	23			14	8
Level 2: From \$20,000 to <\$40,000	19	14			28	22			43	24
Level 3: From \$40,000 to <\$60,000	15	11			49	39			60	34
Level 4: From \$60,000 to <\$80,000	22	16			13	10			34	19
Level 5: From \$80,000 to <\$100,000	23	17			4	3			18	10
Level 6: \$100,000 or more	52	39			4	3			10	6
Total with data	134				127				179	
Mother's education—Highest level completed										
Mean level			4.24 <sub>a</sub>	1.29			3.16 <sub>b</sub>	1.31		
							3.08 <sub>b</sub>	1.28		

 $\chi^2(2, N = 697) = 16.86, p < .001$  $\chi^2(8, N = 684) = 68.42, p < .001$  $\chi^2(6, N = 672) = 79.39, p < .001$  $F(2, 437) = 69.50, p < .001$  $\chi^2(10, N = 440) = 142.04, p < .001$  $F(2, 454) = 32.40, p < .001$

Variable	School A ( <i>n</i> = 212)			School B ( <i>n</i> = 222)			School C ( <i>n</i> = 263)			Statistic
	<i>n</i>	%	<i>M</i>	<i>n</i>	%	<i>M</i>	<i>n</i>	%	<i>M</i>	
1. Some high school	0	0		8	6		8	4		$\chi^2(10, N = 457) = 74.16, p < .001$
2. High school graduate	10	9		47	33		72	36		
3. Some college	23	21		33	23		66	33		
4. College graduate	37	33		34	24		24	12		
5. Some graduate training	12	11		11	8		19	9		
6. Graduate/advanced degree	29	26		10	7		14	7		
Total with data	111			143			203			
Father's education—Highest level completed										
Mean level			4.44 <sub>a</sub>			2.84 <sub>b</sub>			3.00 <sub>b</sub>	$F(2, 435) = 45.84, p < .001$
1. Some high school	1	1		15	11		13	7		$\chi^2(10, N = 438) = 83.99, p < .001$
2. High school graduate	17	16		59	44		80	41		
3. Some college	14	13		22	16		42	21		
4. College graduate	22	20		21	16		32	16		
5. Some graduate training	10	9		5	4		13	7		
6. Graduate or advanced degree	44	41		12	9		16	8		
Total with data	108			134			196			

*Note.* In each row, means with the same subscripts do not significantly differ.

**Table 2**  
Depressive Symptoms by Condition and School: CDI Raw Means (and Standard Deviations)

Assessment point	PRP	PEP	CON	Assessment point	PRP versus CON	PRP versus PEP	PEP versus CON
Schools A+B							
Baseline							
<i>M (SD)</i>	8.57 (7.90)	8.05 (6.99)	8.26 (7.18)				
<i>n</i>	147	144	143				
Post				Post			
<i>M (SD)</i>	6.16 (6.29)	7.09 (7.20)	7.00 (5.69)	ANCOVA	$F(1, 255) = 3.66^{\dagger}$	$F(1, 261) = 3.36^{\dagger}$	$F(1, 255) = 0.01$
<i>n</i>	132	132	126	<i>d</i> (95% CI)	-0.17 (-0.36, 0.01)	-0.18 (-0.37, 0.02)	0.01 (-0.20, 0.21)
6 months				6 months			
<i>M (SD)</i>	5.68 (6.20)	6.12 (5.94)	6.70 (5.29)	ANCOVA	$F(1, 176) = 3.24^{\dagger}$	$F(1, 185) = 1.04$	$F(1, 172) = 0.74$
<i>n</i>	96	92	83	<i>d</i> (95% CI)	-0.24 (-0.50, 0.02)	-0.12 (-0.37, 0.12)	-0.11 (-0.38, 0.15)
12 months				12 months			
<i>M (SD)</i>	4.93 (5.75)	6.01 (7.28)	7.25 (5.90)	ANCOVA	$F(1, 181) = 10.72^{**}$	$F(1, 189) = 2.09$	$F(1, 173) = 2.82^{\dagger}$
<i>n</i>	100	92	84	<i>d</i> (95% CI)	-0.43 (-0.68, -0.17)	-0.19 (-0.44, 0.07)	-0.23 (-0.50, 0.04)
18 months				18 months			
<i>M (SD)</i>	4.98 (6.05)	5.64 (5.72)	6.91 (4.90)	ANCOVA	$F(1, 150) = 9.05^{**}$	$F(1, 155) = 2.43$	$F(1, 144) = 2.66$
<i>n</i>	82	76	71	<i>d</i> (95% CI)	-0.42 (-0.70, -0.15)	-0.21 (-0.47, 0.06)	-0.25 (-0.56, 0.05)
24 months				24 months			
<i>M (SD)</i>	4.40 (5.36)	5.87 (5.32)	7.76 (7.50)	ANCOVA	$F(1, 127) = 8.59^{**}$	$F(1, 132) = 4.21^*$	$F(1, 126) = 1.62$
<i>n</i>	68	67	62	<i>d</i> (95% CI)	-0.47 (-0.78, -0.15)	-0.30 (-0.59, -0.01)	-0.21 (-0.53, 0.12)
30 months				30 months			
<i>M (SD)</i>	4.53 (5.95)	6.70 (7.81)	7.91 (7.91)	ANCOVA	$F(1, 134) = 8.18^{**}$	$F(1, 147) = 4.66^*$	$F(1, 134) = 0.73$
<i>n</i>	75	75	62	<i>d</i> (95% CI)	-0.45 (-0.76, -0.14)	-0.31 (-0.60, -0.03)	-0.14 (-0.47, 0.19)
36 months				36 months			
<i>M (SD)</i>	4.69 (5.91)	7.19 (7.46)	6.18 (6.63)	ANCOVA	$F(1, 121) = 1.94$	$F(1, 125) = 4.25^*$	$F(1, 116) = 0.53$
<i>n</i>	65	63	56	<i>d</i> (95% CI)	-0.23 (-0.56, 0.10)	-0.32 (-0.63, -0.01)	0.11 (-0.23, 0.44)
Baseline				School C			

Assessment point	PRP	PEP	CON	Assessment point	PRP versus CON	PRP versus PEP	PEP versus CON
<i>M</i> ( <i>SD</i> )	8.70 (7.65)	8.22 (6.80)	9.17 (7.58)				
<i>n</i>	85	87	91				
Post				Post			
<i>M</i> ( <i>SD</i> )	8.92 (9.11)	7.33 (7.95)	8.25 (6.57)	ANCOVA	$F(1, 166) = 0.31$	$F(1, 155) = 2.47$	$F(1, 164) = 1.49$
<i>n</i>	80	78	89	<i>d</i> (95% CI)	0.06 (−0.15, 0.27)	0.17 (−0.04, 0.38)	−0.13 (−0.34, 0.08)
6 months				6 months			
<i>M</i> ( <i>SD</i> )	7.90 (7.32)	7.00 (6.87)	7.44 (8.76)	ANCOVA	$F(1, 144) = 0.61$	$F(1, 131) = 0.54$	$F(1, 138) = 0.02$
<i>n</i>	70	64	77	<i>d</i> (95% CI)	0.11 (−0.16, 0.38)	0.09 (−0.16, 0.35)	0.02 (−0.24, 0.28)
12 months				12 months			
<i>M</i> ( <i>SD</i> )	7.26 (6.83)	6.56 (6.95)	7.23 (7.07)	ANCOVA	$F(1, 140) = 0.74$	$F(1, 129) = 1.14$	$F(1, 134) = 0.00$
<i>n</i>	69	63	74	<i>d</i> (95% CI)	0.13 (−0.17, 0.44)	0.15 (−0.13, 0.44)	0.01 (−0.29, 0.30)
18 months				18 months			
<i>M</i> ( <i>SD</i> )	7.53 (7.48)	6.15 (6.65)	7.00 (7.63)	ANCOVA	$F(1, 130) = 1.62$	$F(1, 120) = 2.48$	$F(1, 129) = 0.02$
<i>n</i>	62	61	71	<i>d</i> (95% CI)	0.19 (−0.11, 0.50)	0.23 (−0.06, 0.52)	−0.02 (−0.31, 0.27)
24 months				24 months			
<i>M</i> ( <i>SD</i> )	8.30 (8.08)	5.92 (6.53)	6.97 (7.23)	ANCOVA	$F(1, 113) = 2.06$	$F(1, 105) = 2.18$	$F(1, 115) = 0.00$
<i>n</i>	53	55	63	<i>d</i> (95% CI)	0.22 (−0.09, 0.54)	0.24 (−0.08, 0.56)	0.00 (−0.30, 0.30)
30 months				30 months			
<i>M</i> ( <i>SD</i> )	7.63 (8.50)	5.17 (5.28)	6.53 (7.12)	ANCOVA	$F(1, 87) = 0.93$	$F(1, 82) = 1.67$	$F(1, 86) = 0.09$
<i>n</i>	43	42	47	<i>d</i> (95% CI)	0.19 (−0.20, 0.58)	0.24 (−0.13, 0.61)	−0.06 (−0.44, 0.33)
36 months				36 months			
<i>M</i> ( <i>SD</i> )	9.16 (8.81)	5.09 (5.93)	6.44 (6.79)	ANCOVA	$F(1, 83) = 2.53$	$F(1, 68) = 4.48^*$	$F(1, 78) = 0.40$
<i>n</i>	38	33	48	<i>d</i> (95% CI)	0.34 (−0.08, 0.76)	0.49 (0.03, 0.96)	−0.14 (−0.58, 0.30)

Note. The 95% confidence intervals (CIs) represent the lower bound followed by the upper bound. CDI = Children's Depression Inventory; PRP = Penn Resiliency Program; PEP = Penn Enhancement Program; CON = control; Post = postintervention; ANCOVA = analysis of covariance.

<sup>†</sup>  $p < .10$ .

\*  $p < .05$ .

\*\*  $p < .01$ .